1653

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/645,835A

DATE: 03/16/2001
TIME: 16:06:50

ENTERED

Input Set : A:\Sp36homo.app
Output Set: N:\CRF3\03162001\I645835A.raw

3 <110> APPLICANT: Heinrichs, Jon
4   Johnson, Leslie S.
5   Koenig, Scott
6   Adamou, John E.
8 <120> TITLE OF INVENTION: Pneumococcal Protein Homologs and Fragments for
9   Vaccines
11 <130> FILE REFERENCE: 469201-493
C--> 13 <140> CURRENT APPLICATION NUMBER: US/09/645,835A
14 <141> CURRENT FILING DATE: 2000-08-25
16 <150> PRIOR APPLICATION NUMBER: U.S. 60/150,750
17 <151> PRIOR FILING DATE: 1999-08-25
19 <160> NUMBER OF SEQ ID NOS: 7
21 <170> SOFTWARE: PatentIn Ver. 2.1
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 2478
25 <212> TYPE: DNA
26 <213> ORGANISM: Streptococcus pyogenes
28 <400> SEQUENCE: 1
29 gtgaagaaaa catatggtta tatcggctca gttgctgcta ttttactagc tactcatatt 60
30 ggaagttacc aacttggtaa gcatcatatg ggttcagcaa caaaggacaa tcaaattgcc 120
31 tatattgatg atagcaaagg taaggcaaaa gcccctaaaa caaacaaaac gatggatcaa 180
32 atcagtgctg aagaaggcat ctctgctgaa cagatcgtag tcaaaattac tgaccaaggc 240
33 tatgtgacct cacatggtga ccattatcat ttttacaatg ggaaagttcc ttatgatgcg 300
34 attattagtg aagagttgtt gatgacggat cctaattacc gttttaaaca atcagacgtt 360
35 atcaatgaaa tcttagacgg ttacgttatt aaagtcaatg gcaactatta tgtttacctc 420
36 aagccaggta gcaagcgcaa aaacattcga accaaacaac aaattgctga gcaagtagcc 480
37 aaaggaacta aagaagctaa agaaaaaggt ttagctcaag tggcccatct cagtaaagaa 540
38 gaagttgcgg cagtcaatga agcaaaaaga caaggacgct atactacaga cgatggctat 600
39 atttttagtc cgacagatat cattgatgat ttaggaagatg cttatttagt acctcatggt 660
40 aatcactatc attatattcc taaaaaggat ttgtctccaa gtgagctagc tgctgcacaa 720
41 gcctactgga gtcaaaaaca aggtcgaggt gctagaccgt ctgattaccg cccgacacca 780
42 gccccagccc caggtcgtag gaaagcccca attcctgatg tgacgcctaa tgatgcgtca 840
43 ggtcatcagc cagataacgg tggctatcat ccagcgcctc ctaggccaaa tttagatcaa 900
44 caaaacaaac accaaagaga tgagtttaaa ggaaaaacct ttaaggaact atggtttgat 960
45 ctacaccgtc ttgatttgaa ataccgtcat gtggaagaag atggttgat ttttgaaccg 1020
46 actcaagtga tcaaatcaaa cgcttttggg tatgtggtgc ctcatggaga tcattatcat 1080
47 attatcccaa gaagtcagtt atcacctctt gaaatggaat tagcagatcg atacttagcc 1140
48 ggccaaactg aggacgatga ctcaggttca gatcactcaa aaggcttaga tggtaaacca 1200
49 acacatacct ttcttggtca tcgcatcaaa gcttacggaa ttcattcagt ggataaatca 1260
50 tatgatacga gtgatgctta tgttttttagt aaagaatcca cactatatag gatttggaga 1320
51 ggagttacag ctaaacacgg agatcatttc cactatatag gtcaagctga tgagcttgct 1380
52 tatgagttgg atgaggtcgc taactgggtg aaagcaaaag aaaccactct ttgacactaa 1440
53 gctgctttgg atcaggaaca aggcaaagaa tatatgatgc caaaagatgg caaggactat 1500
54 cgcaaagtaa caaaagatgg taaagtgggc cagattgcct ttgccgaaca agaactaatg 1560
55 ttctatgctc gtgatcaact tgatttgact attgttgaca caggtattga gccacgactt 1620
56 cttaaagata agaaacatta ccgttatgac gctggtaatg ctacttacga tactggaagt 1680
57 gctgtagatg tgtcaagtct gccgatgcat gctggtaatg ctacttacga tactggaagt 1740

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/645,835A

DATE: 03/16/2001
TIME: 16:06:50

Input Set : A:\Sp36homo.app
Output Set: N:\CRF3\03162001\I645835A.raw

```
                                                            gttgacgcgc 1800
58 tcgtttgtta tccctcatat tgatcatatc catgtcgttc cgtattcatg gttgacgcgc 1800
59 gatcagattg caacaatcaa gtatgtgatg caacaccccg aagttcgtcc ggatatatgg 1860
60 tctaagccag ggcatgaaga gtcaggttcg gtcattccaa atgttacgcc tcttgataaa 1920
61 cgtgctggta tgccaaactg gcaaattatc cattctgctg aagaagttca aaaagcccta 1980
62 gcagaaggtc gttttgcaac accagacggc tatattttcg atccacgaga tgttttggcc 2040
63 aaagaaactt ttgtatggaa agatggctcc tttagcatcc caagagcaga tggcagttca 2100
64 ttgagaacca ttaataaatc tgatctatcc caagctgagt ggcaacaagc tcaagagtta 2160
65 ttggcaaaga aaaacgctgg tgatgctact gatacggata aacccaaaga aaagcaacag 2220
66 gcagataaga gcaatgaaaa ccaacagcca agtgaagcca gtaaagaaga agaaaaagaa 2280
67 tcagatgact ttatagacag tttaccagac tatggtctag atagagcaac cctagaagat 2340
68 catatcaatc aattagcaca aaaagctaat atcgatccta agtatctcat tttccaacca 2400
69 gaaggtgtcc aattttataa taaaaatggt gaattggtaa cttatgatat caagacactt 2460
                                                                      2478
70 caacaaataa acccttaa
```

73 <210> SEQ ID NO: 2
74 <211> LENGTH: 825
75 <212> TYPE: PRT
76 <213> ORGANISM: Streptococcus pyogenes

78 <400> SEQUENCE: 2

```
79 Val Lys Lys Thr Tyr Gly Tyr Ile Gly Ser Val Ala Ala Ile Leu Leu
80  1               5                   10                  15
82 Ala Thr His Ile Gly Ser Tyr Gln Leu Gly Lys His His Met Gly Ser
83          20                  25                  30
85 Ala Thr Lys Asp Asn Gln Ile Ala Tyr Ile Asp Asp Ser Lys Gly Lys
86          35                  40                  45
88 Ala Lys Ala Pro Lys Thr Asn Lys Thr Met Asp Gln Ile Ser Ala Glu
89      50                  55                  60
91 Glu Gly Ile Ser Ala Glu Gln Ile Val Val Lys Ile Thr Asp Gln Gly
92  65                  70                  75                  80
94 Tyr Val Thr Ser His Gly Asp His Tyr His Phe Tyr Asn Gly Lys Val
95          85                  90                  95
97 Pro Tyr Asp Ala Ile Ile Ser Glu Glu Leu Leu Met Thr Asp Pro Asn
98          100                 105                 110
100 Tyr Arg Phe Lys Gln Ser Asp Val Ile Asn Glu Ile Leu Asp Gly Tyr
101         115                 120                 125
103 Val Ile Lys Val Asn Gly Asn Tyr Tyr Val Tyr Leu Lys Pro Gly Ser
104         130                 135                 140
106 Lys Arg Lys Asn Ile Arg Thr Lys Gln Gln Ile Ala Glu Gln Val Ala
107 145                 150                 155                 160
109 Lys Gly Thr Lys Glu Ala Lys Glu Lys Gly Leu Ala Gln Val Ala His
110         165                 170                 175
112 Leu Ser Lys Glu Glu Val Ala Ala Val Asn Glu Ala Lys Arg Gln Gly
113         180                 185                 190
115 Arg Tyr Thr Thr Asp Asp Gly Tyr Ile Phe Ser Pro Thr Asp Ile Ile
116         195                 200                 205
118 Asp Asp Leu Gly Asp Ala Tyr Leu Val Pro His Gly Asn His Tyr His
119 210                 215                 220
121 Tyr Ile Pro Lys Lys Asp Leu Ser Pro Ser Glu Leu Ala Ala Ala Gln
122 225                 230                 235                 240
124 Ala Tyr Trp Ser Gln Lys Gln Gly Arg Gly Ala Arg Pro Ser Asp Tyr
```

```
                                                  250                          255
                 245                     Arg Arg Lys Ala Pro Ile Pro
125
127 Arg Pro Thr Pro Ala Pro Ala Pro Gly     265                      270
128               260                 Gly His Gln Pro Asp Asn Gly Gly
130 Asp Val Thr Pro Asn Pro Gly Gln     280                     285
131            275                  Pro Asn Asp Ala Ser Gln Asn Lys His
133 Tyr His Pro Ala Pro Pro Arg     295                     300
134         290                Thr Phe Lys Glu Leu Leu Asp Gln
136 Gln Arg Asp Glu Phe Lys Gly Lys     315                          320
137 305                310            His Val Glu Glu Asp Gly Leu
139 Leu His Arg Leu Asp Leu Lys Tyr Arg     330                       335
140              325               Ser Asn Ala Phe Gly Tyr Val
142 Ile Phe Glu Pro Thr Gln Val Ile Lys     345                     350
143            340                Ile Ile Pro Arg Ser Gln Leu Ser
145 Val Pro His Gly Asp His Tyr His     365
146          355               360  Leu Ala Gly Gln Thr Glu
148 Pro Leu Glu Met Glu Leu Ala Asp Arg Tyr     380
149      370               375             Pro Ser Asp Lys Glu Val
151 Asp Asp Asp Ser Gly Ser Asp His Ser Lys     395                        400
152 385            390                Ala Tyr Gly Lys Gly Leu
154 Thr His Thr Phe Leu Gly His Arg Ile Lys     410                     415
155        405               Ala Tyr Val Phe Ser Lys Glu
157 Asp Gly Lys Pro Tyr Asp Thr Ser Asp     425                     430
158        420               Gly Val Thr Ala Lys His Gly Asp
160 Ser Ile His Ser Val Asp Lys Ser     445
161       435               440   Glu Leu Glu Gln Tyr Glu Leu Asp
163 His Phe His Tyr Ile Gly Phe Gly     460
164      450               455    Gly Gln Ala Asp Glu Leu Ala
166 Glu Val Ala Asn Trp Val Lys Ala Lys     475                          480
167 465           470                Glu Lys Pro Leu Phe Asp Thr
169 Ala Ala Leu Asp Gln Glu Gln Gly Lys     490                     495
170         485               Asp Gly Lys Val Gly Tyr Met
172 Lys Lys Val Ser Arg Lys Val Thr Lys     505                     510
173        500               Tyr Phe Tyr Ala Arg Asp Gln Leu Asp
175 Met Pro Lys Asp Gly Lys Asp Tyr     525
176       515               520    Glu Gln Glu Leu Met Leu Lys Asp Lys
178 Leu Thr Gln Ile Ala Phe Ala Glu     540
179      530               535    Gly Ile Glu Pro Arg Leu
181 Lys His Tyr Arg Tyr Asp Ile Val Asp Thr     555                          560
182 545           550                Ala Gly Asn Ala Thr Tyr
184 Ala Val Asp Val Ser Ser Leu Pro Met His     570                     575
185         565               His Ile Asp His Ile His Val
187 Asp Thr Gly Ser Ser Phe Val Ile Pro     585                          590
188         580               Gln Ile Ala Thr Ile Lys Tyr
190 Val Pro Tyr Ser Trp Leu Thr Arg Asp     600                     605
191        595               Pro Asp Ile Trp Ser Lys Pro Gly
193 Val Met Gln His Pro Glu Val Arg     620
194      610               615    Thr Pro Leu Asp Lys
196 His Glu Glu Ser Gly Ser Val Ile Pro Asn Val     635                640
197 625               630
```

```
199 Arg Ala Gly Met Pro Asn Trp Gln Ile Ile His Ser Ala Glu Glu Val
200                                     650                      655
202 Gln Lys Ala Leu Ala Glu Gly Arg Phe Ala Thr Pro Asp Gly Tyr Ile
203                 660                 665                 670
205 Phe Asp Pro Arg Asp Val Leu Ala Lys Glu Thr Phe Val Trp Lys Asp
206             675                 680                 685
208 Gly Ser Phe Ser Ile Pro Arg Ala Asp Gly Ser Ser Leu Arg Thr Ile
209     690                 695                 700
211 Asn Lys Ser Asp Leu Ser Gln Ala Glu Trp Gln Gln Ala Gln Glu Leu
212 705                 710                 715                 720
214 Leu Ala Lys Lys Asn Ala Gly Asp Ala Thr Asp Thr Asp Lys Pro Lys
215                 725                 730                 735
217 Glu Lys Gln Gln Ala Asp Lys Ser Asn Glu Asn Gln Gln Pro Ser Glu
218             740                 745                 750
220 Ala Ser Lys Glu Glu Glu Lys Glu Ser Asp Asp Phe Ile Asp Ser Leu
221     755                 760                 765
223 Pro Asp Tyr Gly Leu Asp Arg Ala Thr Leu Glu Asp His Ile Asn Gln
224     770                 775                 780
226 Leu Ala Gln Lys Ala Asn Ile Asp Pro Lys Tyr Leu Ile Phe Gln Pro
227 785                 790                 795                 800
229 Glu Gly Val Gln Phe Tyr Asn Lys Asn Gly Glu Leu Val Thr Tyr Asp
230             805                 810                 815
232 Ile Lys Thr Leu Gln Gln Ile Asn Pro
233             820                 825
236 <210> SEQ ID NO: 3
237 <211> LENGTH: 2379
238 <212> TYPE: DNA
239 <213> ORGANISM: Streptococcus pyogenes
241 <400> SEQUENCE: 3
242 atgaaaacga aaaaagttat tattttagtt ggtctattgt tatcatctca gttgactttg 60
243 atagcttgtc aatcacgagg taatggtaca tatcccatta aaacgaaaca atcacgtaag 120
244 ggaatgacgt caaacaaaat taaaccgatt aaaaaaagca aaaagacaaa caagagctcac 180
245 aaaggtgtgg cgggtgtcga ttttcctaca gatgatgacc atgatggtca ttcgcatttt 240
246 aaaatcttat caaaaacaga tcagggaatc gttgttgacc ttattccaaa aggagcaagt 300
247 attttttatg ccgatttaaa gggaagtcca tttgaatacc ttattccaaa aggagcaagt 360
248 ttagctaagc cagctgttgc tcagcgagca gctagtcaag ggacttctaa agtagcagat 420
249 cctcatcacc attatgaatt taacccagcg gatattgtgt ctgaagatgc tttaggctac 480
250 acggttcgcc acgatgatca cttccattat attttgaagt caagcttatc aggtcagaca 540
251 caggcacaag ctaaacaggt tgctactcgc ttgccacaaa ccagtagcct tgtttcaaca 600
252 gctacagcta atggtattcc aggcttgcat ttcccaacct cagatggttt tcaatttaac 660
253 ggtcaaggta ttgttggggt aacaaaagac agtattttag tggaccacga tggtcactta 720
254 catcctattt cttttgcgga ccttcgtcag ggtggctggg cacatgtggc agatcaatac 780
255 gatcccgcta aaaagcaga aaagccagca gaaacccatc agacaccaga gctatctgaa 840
256 cgtgaaaagg aataccaaga aaaattagct tatttggcag aaaaattggg gattgatcca 900
257 tcaactatta aacgtgtgga aacacaagac ggtaaacttg gtttggaata ccctcaccat 960
258 gaccacgcac acgtattgat gttatctgat attgaaatcg gaaaagacat tccagatcca 1020
259 catgctattg agcatgcccg tgaattggaa aaacataagg ttggaatgga taccttgcgt 1080
260 gcctagggt ttgatgaaga agtgattttg gatatcgttc gcactcacga tgctccaacc 1140
261 ccattcccat caaatgaaaa agatccgaat atgatgaaag aatggttagc aacggttatc 1200
```

```
                                                     actgttaccc 1260
262 aaacttgact tgggcagccg taaagatcct ttgcaacgta aaggactttc tgttttgcaa 1320
263 aacttagaaa ctttaggaat tggctttaca ccaatcaaag atatctcacc tagattttg 1380
264 tttaaaaaat tgaaacagtt gttaatgaca aaaacagggg tgactgatta acaatctcaa agatattagt 1440
265 gataatatgc cacagttaga aggcattgat atttcacaaa tattgagtaa caataagatt 1500
266 ttcttgagca aatatataaaa cttaactcta gtagcggctg tattgagtaa caataagatt 1560
267 attaggccgc ttggtcaatt accaaatctc aaattcctcg aattgcacat tgataataat 1620
268 tctgatttaa gcccactggc atcgttacat caattgcaag aattgcacat tgatttatca 1680
269 cagattacag atttaagccc tgtttctcat aaagaatcat tgacggttgt tgatttatca 1740
270 agaaatgctg atgttgactt agcaacactt caagcaccca aattagaaac gttaatggtc 1800
271 aatgatacca aggtttctca tttggatttc ttgaaaaata atcctaatct atctagccta 1860
272 tctattaacc gtgcgcaatt gcaatctctt gaaggtattg aagcaagtag cgtcattgtc 1920
273 agagtagaag cagaaggtaa ccaaattaaa tcgcttgtgc ttaaagacaa gcaagggtca 1980
274 cttactttct tggatgtgac cgtgtctaaa aaccaattaa caaatgtcaa cctatctaaa 2040
275 acagcacttg acattttaag cattgatatt agtcataaca atatctcatt agcagacctt 2100
276 cccaataaga cagttactaa tccagaagcc attgcgaaaa acttcccagc ggtttacgaa 2160
277 aaattgaacg agcaacatat tccagaagcc attgcgaaaa acttcccagc ggtttacgaa 2220
278 ggttctatgg taggtaatgg aacagctgaa gaaaaagcag ctatggctac taaggcgaaa 2280
279 gaaagtgctc aagaagcatc ggaatcacat gactacaacc ataatcatac ctatgaagat 2340
280 gaagaaggtc atgctcacga gcacagagac aaagatgatc acgaccatga acatgaggat 2379
281 gaaaatgaag ctaaagatga gcaaaaccat gctgactaa
```

```
284 <210> SEQ ID NO: 4
285 <211> LENGTH: 792
286 <212> TYPE: PRT
287 <213> ORGANISM: Streptococcus pyogenes
289 <400> SEQUENCE: 4
290 Met Lys Thr Lys Lys Val Ile Ile Leu Val Gly Leu Leu Leu Ser Ser
291  1               5                   10                  15
293 Gln Leu Thr Leu Ile Ala Cys Gln Ser Arg Gly Asn Gly Thr Tyr Pro
294          20                  25                  30
296 Ile Lys Thr Lys Gln Ser Arg Lys Gly Met Thr Ser Asn Lys Ile Lys
297          35                  40                  45
299 Pro Ile Lys Lys Ser Lys Lys Thr Asn Lys Thr His Lys Gly Val Ala
300      50                  55                  60
302 Gly Val Asp Phe Pro Thr Asp Asp Gly Phe Ile Leu Thr Lys Asp Ser
303  65                  70                  75                  80
305 Lys Ile Leu Ser Lys Thr Asp Gln Gly Ile Val Val Asp His Asp Gly
306              85                  90                  95
308 His Ser His Phe Ile Phe Tyr Ala Asp Leu Lys Gly Ser Pro Phe Glu
309              100                 105                 110
311 Tyr Leu Ile Pro Lys Gly Ala Ser Leu Ala Lys Pro Ala Val Ala Gln
312          115                 120                 125
314 Arg Ala Ala Ser Gln Gly Thr Ser Lys Val Ala Asp Pro His His His
315      130                 135                 140
317 Tyr Glu Phe Asn Pro Ala Asp Ile Val Ala Glu Asp Ala Leu Gly Tyr
318  145                 150                 155                 160
320 Thr Val Arg His Asp Asp His Phe His Tyr Ile Leu Lys Ser Ser Leu
321              165                 170                 175
323 Ser Gly Gln Thr Gln Ala Gln Ala Lys Gln Val Ala Thr Arg Leu Pro
324              180                 185                 190
```

3/16/01

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/645,835A

DATE: 03/16/2001
TIME: 16:06:51

Input Set : A:\Sp36homo.app
Output Set: N:\CRF3\03162001\I645835A.raw

L:13 M:270 C: Current Application Number differs, Replaced Current Application Number